Image Deconvolution by Local Order Preservation of Pixels Values

Stéphanie Guérit*, Laurent Jacques*†, John A. Lee*†‡
*ICTEAM/ELEN, ISPGroup, ‡MIRO/IREC, Université catholique de Louvain, Belgium

***Abstract*—Positron emission tomography is more and more used in radiation oncology, since it conveys useful functional information about cancerous lesions. Its rather low spatial resolution, however, prevents accurate tumor delineation and heterogeneity assessment. Post-reconstruction deconvolution with the measured point-spread function can address this issue, provided it does not introduce undesired artifacts. These usually result from inappropriate regularization, which is either absent or making too strong assumptions about the structure of the signal. This paper proposes a deconvolution method that is based on inverse problem theory and involves a new regularization term that preserves local pixel value order relationships. Such regularization entails relatively mild constraints that are directly inferred from the observed data. This paper investigates the theoretical properties of the proposed regularization and describes its numerical implementation with a primal-dual algorithm. Preliminary experiments with synthetic images are presented to compare quantitatively and qualitatively the proposed method to other regularization schemes, like TV and TGV.**

I. Introduction

Image restoration methods often address ill-posed problems and, without proper regularization, solving these can converge to implausible images, with undesired artifacts. However, common regularization schemes typically entail some strong *a priori* about the image content, which can also favor the generation of other artifacts. Good solutions must hence reach a tradeoff that mitigates adequately both trends. In the context of medical imaging, non-uniqueness of the solution, artifacts, and tedious parameter tuning cause defiance and make potential users reluctant to use image restoration.

This paper finds its motivation in the use of positron emission tomography (PET) in radiation oncology (RO). PET conveys useful functional information about tumors but suffers from rather low spatial resolution, while RO, interpretable as a ballistic problem, crucially depends on their accurate delineation [1], [2]. Based on the knowledge of the PET system point-spread function (PSF), image deconvolution is feasible. However, its adoption remains rather limited, either as a quantitative restoration tool [3] or as preprocessing to improve image segmentation [4]–[6]. Similarly, deconvolution integrated within PET image reconstruction and modeled by accounting for blur in the forward model of the PET system remains a challenging issue in both academia [7], [8] and industry [9].

The driving idea of this paper relies on two assumptions. (i) Large intensity variations and discontinuities in PET images are mostly distant from each other, on the scale of the PSF width. In other words, intensity profiles along each dimension could be split into only a few monotonic segments. Textured patterns, like checkerboards or stripes, would then be uncommon, except in images of phantoms specifically designed to measure spatial resolution (like flanged Jaszczak phantoms with collections of cold rods). (ii) Convolving some ground truth image with the PSF has little impact on local order relationships between the value of neighboring pixels. Well-regularized deconvolution of the blurred image with the same PSF should hence also affect weakly these relationships. Non-regularized deconvolution, though, is known to amplify some high frequencies, which typically causes Gibbs ringing artifacts to appear. Such spurious oscillations, on the other hand, clearly violate local order constraints on pixel values.

This paper proposes a deconvolution method with a regularization scheme that enforces the preservation of local order relationships. In practice, this is achieved by constraining the sign of the derivatives approximated with finite differences. Preventing such sign changes is obviously milder than other regularization schemes that favour sparsity of finite differences, like total variation (TV) [10]. Those amount to assuming that gradient norm is null almost everywhere, which yields potentially cartoon-like or blocky images with staircasing artifacts. In addition to formulating the inverse problem of deconvolution with local order preservation and solving it numerically, the main contributions of this paper are also to investigate the validity of local order preservation and how to estimate order relationships from noisy data.

The rest of this paper is organised as follows. Sec. II introduces the forward model we assume for PET images. Sec. III formulates the inverse problem with local order-preserving constraints. Sec. IV investigates the validity of local order preservation and how order relationships can be identified in spite of noise. Sec. V details the numerical implementation. Sec. VI and Sec. VII present the experiment setup and discuss the results. Eventually, Sec. VIII draws conclusions and sketches perspectives for future work.

II. Forward model

For the purpose of simplicity, we directly adopt a discrete framework. The vectorized d-dimensional signals of size $N_1 \times \cdots \times N_d$ belong to the Euclidean space $\mathbb{R}^N$ with $N = \prod_{i=1}^{d} N_i$. Let $\boldsymbol{u}_0 \in \mathbb{R}^N$ be the original (unknown) PET image of the radioactivity in a patient body. After tomographic reconstruction, we observe $\boldsymbol{z} \in \mathbb{R}^N$, a version of $\boldsymbol{u}_0$ blurred by operator $\boldsymbol{H}$ and corrupted by a noise $\mathcal{N}_\nu$ of parameter ν,

$$\boldsymbol{z} = \mathcal{N}_\nu(\boldsymbol{H}(\boldsymbol{u}_0)).$$

$\boldsymbol{H}$ accounts for both the physical and instrumental inaccuracies limiting the spatial resolution of the acquired data. In the

† Research Associates with the Belgian F.R.S.-FNRS.

978-0-9928-6265-7/16/$31.00 ©2016 IEEE

2D-plane, it is usually measured by imaging a point source of radioactivity [11]. We assume in first approximation that the blur is uniform ($\boldsymbol{H}$ is linear), Gaussian and isotropic near the center of the field-of-view [11]. The blurred image therefore results from the convolution of $\boldsymbol{u}_0$ with $\boldsymbol{h} \in \mathbb{R}^N$, the kernel associated with $\boldsymbol{H} \in \mathbb{R}^{N\times N}$. We consider the noise as additive and Gaussian: $\boldsymbol{n} \in \mathbb{R}^N$ with $n_i \underset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma_n^2)$, $i \in \{1, \dots, N\}$. Then,

$$\boldsymbol{z} = \boldsymbol{H}\boldsymbol{u}_0 + \boldsymbol{n}. \tag{1}$$

Regarding the physics of PET imaging, we constraint model (1) with (i) *positivity* of $\boldsymbol{u}_0$ (measure of a nonnegative radioactivity) and (ii) *photometry invariance*, *i.e.*, the preservation of total photon counts in the original and observed images.

III. Inverse problem formulation

From observations $\boldsymbol{z}$, we want to find the best estimate $\hat{\boldsymbol{u}}_0 \in \mathbb{R}^N$ of $\boldsymbol{u}_0$, considering the impact of the noise and such that $\boldsymbol{H}\hat{\boldsymbol{u}}_0 \simeq \boldsymbol{z}$. Finding such estimate is challenging: the problem could be ill-posed, *e.g.*, if $\boldsymbol{H}$ is ill-conditioned and consequently, non invertible. We have no guarantee on the unicity of the solution. To fix this issue and reduce the set of feasible solutions, a widely used method is to regularize the problem [12]. A penalty term encourages the solution to respect a certain prior model of the original image (*e.g.*, sparsity in specific basis, small TV-norm, etc. [12], [13]). These *a priori* assumptions are often valid for particular classes of signals and do not directly depend on the observations.

A. Local constraints

The key idea of our method is to regularize the problem by locally constraining the order relationships between the value of neighboring signal samples. The constraints are inferred from the observed data $\boldsymbol{z}$ and in this way, do not assume a direct prior model on the signal $\boldsymbol{u}_0$.

Definition 1. *The discrete d-dimensional gradient operator ∇ applicable to signals in $\mathbb{R}^N$ is defined as*

$$\nabla : \mathbb{R}^N \to \mathbb{R}^{dN}, \boldsymbol{x} \mapsto (\nabla\boldsymbol{x}) = (\nabla_1\boldsymbol{x}^T, \dots, \nabla_d\boldsymbol{x}^T)^T,$$

with $\nabla_i \in \mathbb{R}^{N\times N}$ the first spatial derivative in direction $\boldsymbol{e}_i$, approximated using forward first-order finite differences.

The matrix $\boldsymbol{A} \in \mathbb{R}^{dN\times N}$ is sparse and defines the order constraints between the value of neighboring elements of $\boldsymbol{u}_0$. Ideally, if $\boldsymbol{u}_0$ was known at restoration, it is defined by $\boldsymbol{A} \coloneqq -\text{diag}(\text{sign}(\nabla\boldsymbol{u}_0))\nabla$, so that $\boldsymbol{u} = \boldsymbol{u}_0$ satisfies $\boldsymbol{A}\boldsymbol{u} \preceq \boldsymbol{0}$.

B. Strict and relaxed formulations

The non-smooth convex problem is

$$\begin{aligned} \underset{\boldsymbol{u}\in\mathbb{R}^N}{\text{minimize}} \quad & \|\boldsymbol{H}\boldsymbol{u} - \boldsymbol{z}\|_2^2 \\ \text{subject to} \quad & \boldsymbol{A}\boldsymbol{u} \preceq \boldsymbol{0},\ \boldsymbol{u} \succeq \boldsymbol{0},\ \textstyle\sum_{i=1}^N u_i = \sum_{i=1}^N z_i, \end{aligned} \tag{2}$$

where the last constraint results from photometry invariance property. A relaxed formulation of problem (2) could tolerate the violation of some constraints, as long as they do not accumulate "too much". The constraint $\boldsymbol{A}\boldsymbol{u} \preceq \boldsymbol{0}$ is replaced with a one-sided ℓ_1-norm defined as $\|(\boldsymbol{x})_+\|_1 = \|\max(\boldsymbol{0}, \boldsymbol{x})\|_1$,

$$\begin{aligned} \underset{\boldsymbol{u}\in\mathbb{R}^N}{\text{minimize}} \quad & \|(\boldsymbol{A}\boldsymbol{u})_+\|_1 \\ \text{subject to} \quad & \boldsymbol{H}\boldsymbol{u} \in \mathcal{B},\ \boldsymbol{u} \succeq \boldsymbol{0},\ \textstyle\sum_{i=1}^N u_i = \sum_{i=1}^N z_i, \end{aligned} \tag{3}$$

where $\mathcal{B} = \{\boldsymbol{x} \in \mathbb{R}^N | \|\boldsymbol{x} - \boldsymbol{z}\|_2^2 \le \sigma_n^2(N + c\sqrt{N}), c = \mathcal{O}(1)\}$ since we have some knowledge on the statistics of the noise.

C. Approximation of matrix A

Assuming that $\boldsymbol{H}$ preserves most of the order relationships of $\boldsymbol{u}_0$, we approximate $\boldsymbol{A}$ with $\boldsymbol{A}_{\boldsymbol{H}} \coloneqq -\text{diag}(\text{sign}(\nabla\boldsymbol{H}\boldsymbol{u}_0))\nabla$. $\boldsymbol{H}$ is said to be *order-preserving* and we can bound the number of wrong inequalities in $\boldsymbol{A}_{\boldsymbol{H}}$ with respect to $\boldsymbol{A}$ (see Def. 3).

Definition 2. *The normalized Hamming distance between $\boldsymbol{x}$ and $\boldsymbol{y} \in \mathbb{R}^N$ is defined as $d_H(\boldsymbol{x}, \boldsymbol{y}) = \frac{1}{N}\sum_{i=1}^N x_i \oplus y_i$, where symbol $\oplus$ denotes the logical XOR operation.*

Definition 3. *Operator $\boldsymbol{H}$ is* order-preserving *with parameter $\tau \in [0, 1]$ on the class of functions $\mathcal{A}$ if*

$$d_H(\mathit{sign}(\nabla\boldsymbol{H}\boldsymbol{x}), \mathit{sign}(\nabla\boldsymbol{x})) \le \tau, \qquad \forall\boldsymbol{x} \in \mathcal{A},$$

subsequently denoted by OP$(\tau, \mathcal{A})$.

When observations are corrupted only by a blurring operator $\boldsymbol{H}$, *i.e.*, $\boldsymbol{z} = \boldsymbol{H}\boldsymbol{u}_0$, then $\boldsymbol{A}_{\boldsymbol{H}} \coloneqq -\text{diag}(\text{sign}(\nabla\boldsymbol{z}))\nabla$. When data are corrupted by both blur and noise, *i.e.*, $\boldsymbol{z} = \boldsymbol{H}\boldsymbol{u}_0 + \boldsymbol{n}$, order relationships are mostly not preserved in $\boldsymbol{z}$ anymore. From observations $\boldsymbol{z}$, we estimate the gradient of the noiseless (but still blurred) observations $\boldsymbol{v}_0 \coloneqq \nabla\boldsymbol{H}\boldsymbol{u}_0$, by solving

$$\underset{\boldsymbol{v}\in\mathbb{R}^{dN}}{\text{minimize}}\ \|(\nabla, \dots d\text{-times} \dots, \nabla)\boldsymbol{v}\|_{2,1} + \imath_{\mathcal{D}}(\boldsymbol{v}),$$

where $\mathcal{D} = \{\boldsymbol{x} \in \mathbb{R}^N | \|\boldsymbol{x} - \nabla\boldsymbol{z}\|_2^2 \le 2\sigma_n^2(N + c\sqrt{N})\}$ with $c = \mathcal{O}(1)$ [14], [15]. The TV prior prevents oscillations of $\boldsymbol{v}$.

IV. On order preservation

This section provides upper bounds of the normalized quantity of violated inequalities in $\boldsymbol{A}_{\boldsymbol{H}}$, *i.e.*, $d_H(\text{sign}(\nabla(\boldsymbol{H}\boldsymbol{x} + \boldsymbol{n}), \text{sign}(\nabla\boldsymbol{x}))$. This theoretical analysis characterizes among others the OP property and its parameter τ for a particular class of 1D piecewise monotonic signals.

Definition 4. *$\mathcal{C}(S)$ denotes the class of ℓ^2-functions: $\Omega \coloneqq \{1, \dots, N\} \to [0, 1]$, monotonic on each of their S subdomains Ω_j, with $S \ge 1$ and $j \in \{1, \dots, S\}$.*

Proposition 1. *Let $\boldsymbol{h} \in \mathbb{R}^N$ be a non-negative kernel with $|\mathit{supp}(\boldsymbol{h})| = 2p+1$. Then, the associated convolution operator $\boldsymbol{H} \in \mathbb{R}^{N\times N}$ is OP$(\tau, \mathcal{C}(S))$ with $\tau = \min(2pS/N, 1)$.*

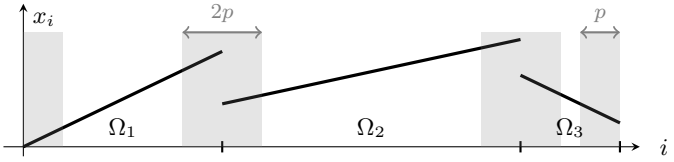

Fig. 1: Signal $\boldsymbol{x} \in \mathcal{C}(3)$ on $\Omega = \Omega_1 \cup \Omega_2 \cup \Omega_3$. Preservation of order relationships is uncertain for elements in gray areas.

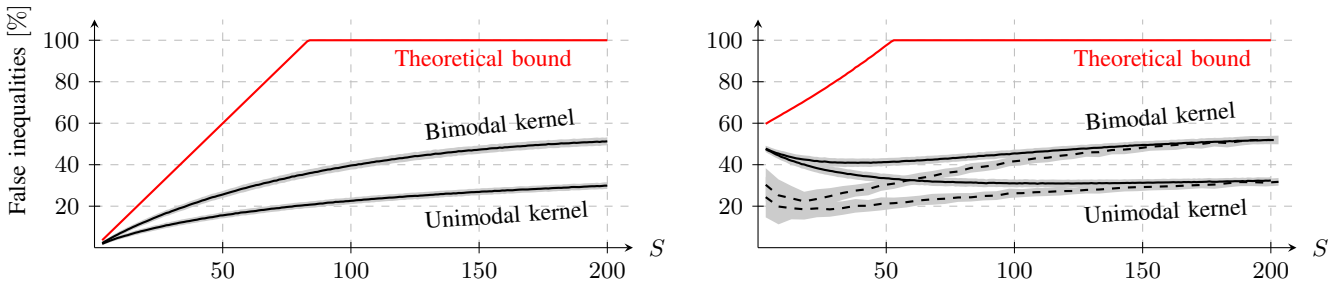

Fig. 2: Mean proportion (1000 realizations) of wrong inequalities in $\boldsymbol{A_H}$ as a function of the number of sub-domains S for 1D piecewise-linear signals in $\mathcal{C}(S)$ ($N = 1001$) blurred by Gaussian unimodal or bimodal kernel ($\sigma = 2$, $|\text{supp}| = 13$) and (right) corrupted by Gaussian noise ($\sigma_n = 0.01$). Solid lines: $\boldsymbol{A_H}$ computed from $\boldsymbol{z}$; dashed lines: $\boldsymbol{A_H}$ computed from $\hat{\boldsymbol{v}}$. Gray areas represent the standard deviation.

Proof. Let $\boldsymbol{x}$ be a signal in $\mathcal{C}(1)$, monotonically increasing on Ω. The discrete convolution with kernel $\boldsymbol{h}$ is defined as

$$y_n = \sum_{m=-p}^{p} h_m x_{n-m}, \qquad \forall n \in \Omega,$$

with unknown boundaries conditions. Let $\mathcal{J} = \{1, \dots, p\} \cup \{N-p+1, \dots, N\}$ be the set of $2p$ elements at the boundaries of the domain. Let $i, j \in \Omega \setminus \mathcal{J}$ such that $i \leq j$ and $x_i \leq x_j$. Then,

$$y_i = \sum_{m=-p}^{p} h_m x_{i-m} \leq y_j = \sum_{m=-p}^{p} h_m x_{j-m},$$

since $h_m \geq 0$ and $x_{i-m} \leq x_{j-m}, \forall m \in \{-p, \dots, p\}$. Operator $\boldsymbol{H}$ preserves the gradient sign of $\boldsymbol{x}$ on $\Omega \setminus \mathcal{J}$. On $\mathcal{J}$, preservation is uncertain and depends on the boundaries conditions. Now, let $\boldsymbol{x}$ be a signal in $\mathcal{C}(S)$. On each sub-domain Ω_j, the signal belongs to $\mathcal{C}(1)$ (see Fig. 1). On $\mathcal{J} = \cup_{j=1}^{S} \mathcal{J}_j$, the preservation of order relationships is uncertain. We can bound the number of non-preserved inequalities by the number of uncertain ones, *i.e.*, the cardinality of $\mathcal{J}$,

$$N d_H(\text{sign}(\nabla \boldsymbol{H}\boldsymbol{x}), \text{sign}(\nabla \boldsymbol{x})) \leq |\mathcal{J}| \quad \in [0, N]\,.$$

In the worst case, $|\mathcal{J}_j| = 2p, \forall j \in \{1, \dots, S\}$, and the normalized number of uncertain relationships is $2pS/N$. □

Let $\boldsymbol{x} \in \mathcal{C}(S)$. Distance $d_H(\text{sign}(\nabla(\boldsymbol{H}\boldsymbol{x} + \boldsymbol{n})), \text{sign}(\nabla \boldsymbol{x}))$ can be rewritten as $d_H(\text{sign}(\nabla \boldsymbol{H}\boldsymbol{x} + \tilde{\boldsymbol{n}}), \text{sign}(\nabla \boldsymbol{x}))$. Since ∇ is approximated by forward first-order finite differences, $\tilde{n}_i = (\nabla \boldsymbol{n})_i \sim \mathcal{N}(0, 2\sigma_n^2)$ and neighbors elements $\tilde{n}_i$ and $\tilde{n}_{i+1}$ are not independent anymore. As Hamming distance satisfies the triangle inequality,

$$\begin{aligned} d_H(\text{sign}(\nabla \boldsymbol{H}\boldsymbol{x} + \tilde{\boldsymbol{n}}), \text{sign}(\nabla \boldsymbol{x})) &\leq d_H(\text{sign}(\nabla \boldsymbol{H}\boldsymbol{x}), \text{sign}(\nabla \boldsymbol{x})) \\ &\quad + d_H(\text{sign}(\nabla \boldsymbol{H}\boldsymbol{x} + \tilde{\boldsymbol{n}}), \text{sign}(\nabla \boldsymbol{H}\boldsymbol{x})). \end{aligned}$$

The first term is deterministic and imputable to the blurring model. It is bounded by τ (see Prop. 1), since $\boldsymbol{h}$ in (1) is assumed to be Gaussian and is consequently non-negative with a finite support $\{-p, \dots, p\}$ including more than 99% of its total energy. The second term, denoted $d_{\tilde{\boldsymbol{n}}}$, can be considered as the stochastic error due to the noise.

By linearity, the expected value of $d_{\tilde{\boldsymbol{n}}}$ is equal to the normalized sum of element expectations E_i bounded by

$$\begin{aligned} E_i &= P\left[((\nabla \boldsymbol{H}\boldsymbol{x})_i + \tilde{n}_i)(\nabla \boldsymbol{H}\boldsymbol{x})_i \leq 0\right] = P(\tilde{n}_i \geq |(\nabla \boldsymbol{H}\boldsymbol{x})_i|) \\ &= \int_{\frac{|(\nabla \boldsymbol{H}\boldsymbol{x})_i|}{\sqrt{2}\sigma_n}}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-s^2/2}\, \mathrm{d}s \leq \frac{1}{2} e^{-(\nabla \boldsymbol{H}\boldsymbol{x})_i^2/4\sigma_n^2}. \end{aligned}$$

The tail of standard normal distribution $\int_t^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-t^2/2}\, \mathrm{d}t$ is bounded by $\frac{1}{2} e^{-t^2/2}$, for $t \geq 0$ [16]. Independence of odd and of even elements in $d_{\tilde{\boldsymbol{n}}} = d_{\tilde{\boldsymbol{n}},\text{odd}} + d_{\tilde{\boldsymbol{n}},\text{even}}$ allows us to use the following Hoeffding inequality [14] and the union bound,

$$\begin{aligned} P\left(d_{\tilde{\boldsymbol{n}}} \geq \tfrac{1}{2N}\left(\textstyle\sum_{\text{odd}} e^{-(\nabla \boldsymbol{H}\boldsymbol{x})_i^2/4\sigma_n^2} + \sum_{\text{even}} e^{-(\nabla \boldsymbol{H}\boldsymbol{x})_i^2/4\sigma_n^2}\right) + 2\delta\right) \\ \leq 2P(d_{\tilde{\boldsymbol{n}},\text{odd}} \geq \tfrac{1}{2N} \textstyle\sum_{\text{odd}} e^{-(\nabla \boldsymbol{H}\boldsymbol{x})_i^2/4\sigma_n^2} + \delta) \\ \leq 2P(d_{\tilde{\boldsymbol{n}},\text{odd}} \geq \tfrac{1}{N} \textstyle\sum_{\text{odd}} E_i + \delta) \leq 2e^{-4N\delta^2}, \end{aligned}$$

with $\delta > 0$, chosen such that $2e^{-4N\delta^2} = 0.01$. Finally,

$$\begin{aligned} &d_H(\text{sign}(\nabla \boldsymbol{H}\boldsymbol{x} + \tilde{\boldsymbol{n}}), \text{sign}(\nabla \boldsymbol{x})) \\ &\leq \min\left(\tfrac{2pS}{N} + \tfrac{1}{2N} \textstyle\sum_{i=1}^{N} e^{-(\nabla \boldsymbol{H}\boldsymbol{x})_i^2/4\sigma_n^2} + 2\sqrt{\tfrac{\ln 200}{4N}}, 1\right), \end{aligned} \quad (4)$$

with probability greater than 0.99. In Fig. 2, the theoretical bounds are compared to the empirical ones obtained for different kernels $\boldsymbol{k}$ and different levels of noise.

V. Numerical implementation

In this section, we introduce the Chambolle-Pock (CP) primal-dual algorithm [13] in order to solve problems (2) and (3), introduced in Sec. III. It belongs to the family of proximal algorithms, that can be used for optimization of nonsmooth and non differentiable objective functions. The CP algorithm is designed to solve the saddle-point problem

$$\underset{\boldsymbol{x} \in \mathcal{X}}{\text{minimize}} \; \underset{\boldsymbol{y} \in \mathcal{Y}}{\text{maximize}} \quad G(\boldsymbol{x}) + \langle \boldsymbol{L}\boldsymbol{x}, \boldsymbol{y} \rangle - F^\star(\boldsymbol{y}), \quad (5)$$

where $\boldsymbol{L} : \mathcal{X} \to \mathcal{Y}$ is a linear continuous operator with a norm $\|\boldsymbol{L}\|_2 := \max\{\|\boldsymbol{L}\boldsymbol{x}\|_2 \mid \boldsymbol{x} \in \mathcal{X} \text{ with } \|\boldsymbol{x}\|_2 \leq 1\}$. $G : \mathcal{X} \to [0, +\infty[$ and $F^\star : \mathcal{Y} \to [0, +\infty[$ are proper, convex and lower-semicontinuous (l.s.c.) functions. $F^\star$ is the Legendre-Fenchel conjugate of F. Formulation (5) is a primal-dual formulation of the primal problem $\text{minimize}_{\boldsymbol{x} \in \mathcal{X}}\; F(\boldsymbol{L}\boldsymbol{x}) + G(\boldsymbol{x})$.

Let φ be a l.s.c. convex function from $\mathcal{X}$ to $]-\infty, +\infty[$ such that domφ is non empty. The proximal operator of $\varphi : \mathcal{X} \to \mathcal{X}$ evaluated in $\tilde{\boldsymbol{x}} \in \mathcal{X}$ is unique and defined as $\text{prox}_{\varphi}(\tilde{\boldsymbol{x}}) := \arg\min_{\boldsymbol{x} \in \mathcal{X}} \frac{1}{2}\|\tilde{\boldsymbol{x}} - \boldsymbol{x}\|_2^2 + \varphi(\boldsymbol{x})$ [17].

A compact formulation of (2) and (3) is given by

$$\underset{\boldsymbol{u} \in \mathbb{R}^N}{\text{minimize}} \quad F_1(\boldsymbol{H}\boldsymbol{u}) + F_2(\boldsymbol{A}\boldsymbol{u}) + G(\boldsymbol{u}),$$

where $F_1(\boldsymbol{x}) = \imath_{\mathcal{B}}(\boldsymbol{x})$ with $\mathcal{B} = \{\boldsymbol{x} \in \mathbb{R}^N \| \|\boldsymbol{x} - \boldsymbol{z}\|_2^2 \leq \sigma_n^2(N + c\sqrt{N}), c = \mathcal{O}(1)\}$. Function $F_2(\boldsymbol{x}) = \imath_{\mathbb{R}_-^N}(\boldsymbol{x})$ in (2)

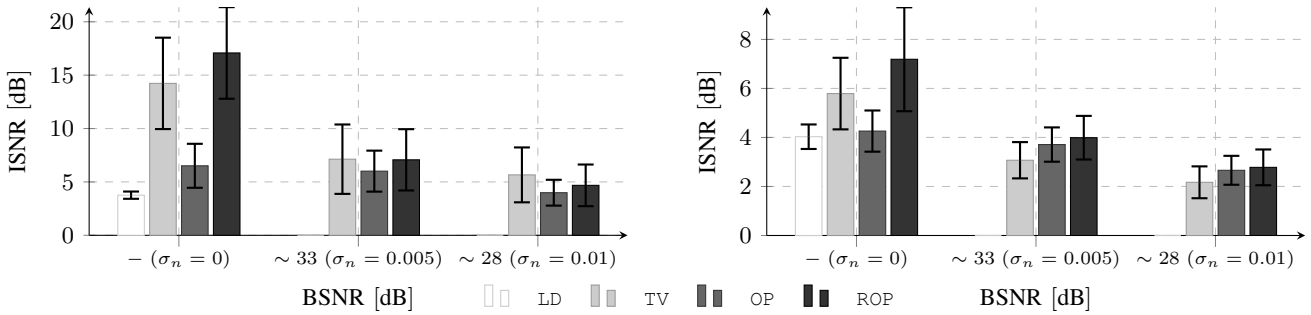

Fig. 3: Mean ISNR (100 realizations) of 1D piecewise-cubic signals of length $N = 1001$ in $\mathcal{C}(S)$ (left: $S = 7$, right: $S = 47$) restored with different methods (LD, TV, OP and ROP). Original signals are blurred by Gaussian unimodal kernel ($\sigma = 2$, $|\text{supp}| = 13$) and corrupted by additive Gaussian noise ($\sigma_n = 0$, $\sigma_n = 0.005$ and $\sigma_n = 0.01$). Error bars represent the standard deviation of the sample.

and $F_2(\boldsymbol{x}) = \|\boldsymbol{x}_+\|_1$ in (3). Finally, $G(\boldsymbol{x}) = \imath_{\mathbb{R}_+^N}(\boldsymbol{x}) + \imath_{\mathcal{P}}(\boldsymbol{x})$ with $\mathcal{P} = \{\boldsymbol{x} \in \mathbb{R}^N | \sum_{i=1}^N x_i = \sum_{i=1}^N z_i\}$.

A primal-dual formulation similar to (5) is derived in a product space [18] using duality principles [12] and leads to Algorithm 1. Let $\tilde{\boldsymbol{x}} \in \mathbb{R}^N$. The proximal operator of the primal function G is given by a combination of positivity (max) and photometry invariance (average) proximal operators [17],

$$\text{prox}_{\sigma G}(\tilde{\boldsymbol{x}}) = \max(\tilde{\boldsymbol{x}}, 0) - \tfrac{1}{N}\sum_{i=1}^N (\max(\tilde{\boldsymbol{x}}, 0)) - \boldsymbol{z})_i \,.$$

Using the Moreau decomposition property [17], the proximal operator of $F_1^\star$ is $\text{prox}_{\sigma^{(n)} F_1^\star}(\tilde{\boldsymbol{x}}) = \tilde{\boldsymbol{x}} - \sigma^{(n)} \Pi_{\mathcal{B}}(\tilde{\boldsymbol{x}}/\sigma^{(n)})$, where $\Pi_{\mathcal{B}}$ is the projection on the ℓ_2-ball $\mathcal{B}$. In formulation (2), the proximal operator of $F_2^\star$ is $\text{prox}_{\sigma^{(n)} F_2^\star}(\tilde{\boldsymbol{x}}) = \max(\tilde{\boldsymbol{x}}, \boldsymbol{0})$. In formulation (3), the proximal operator of $F_2^\star$ is

$$\text{prox}_{\sigma^{(n)} F_2^\star}(\tilde{\boldsymbol{x}}) = \tilde{\boldsymbol{x}} - \min\left(\max\left(|\tilde{\boldsymbol{x}}| - 1, \boldsymbol{0}\right) \cdot \text{sign}(\tilde{\boldsymbol{x}}), \tilde{\boldsymbol{x}}\right).$$

VI. Experiments

Simulations are done first on 1D piecewise-cubic signals of length $N = 1001$ with S sub-domains. Original signals are blurred by normalized Gaussian unimodal or bimodal kernel ($\sigma = 2$) and eventually corrupted by additive Gaussian noise. The bimodal kernel is obtained from the unimodal one by flipping its left and right halves. Experiment 1 computes the mean Hamming distance between $\boldsymbol{A}$ and $\boldsymbol{A_H}$ for different S and compares it to the theoretical bounds of Prop. 1 and (4). Experiment 2 compares the quality of signals restored with different deconvolution methods: Landweber with positivity constraint (LD) [19], TV regularization (TV) [20], local order preservation (OP) and relaxed local order preservation (ROP).

The synthetic image used for 2D experiments is a modified Shepp-Logan phantom of size 128×128, with intensities in $[0, 255]$. The constant surfaces of original phantom have been replaced with a mixture of affine, Gaussian and sinusoidal functions, more representative of real medical data. Original images are convolved with isotropic Gaussian kernel ($\sigma = 1.2$ and $\sigma = 2$) and corrupted by additive Gaussian noise ($\sigma_n = 0.5$, $\sigma_n = 1$ and $\sigma_n = 2.5$). Experiment 3 compares the quality of images restored with TV, total generalized variation regularization (TGV) [10], OP and ROP.

The quality of restored signals $\hat{\boldsymbol{u}}_0$ and observations is measured using the increase in SNR, $\text{ISNR} = 20\log_{10}\|\boldsymbol{z} - \boldsymbol{u}_0\|_2 / \|\hat{\boldsymbol{u}}_0 - \boldsymbol{u}_0\|_2$ and the blurred SNR, $\text{BSNR} = 10\log_{10}\text{var}(\boldsymbol{H}\boldsymbol{u}_0)/\sigma_n^2$, respectively. Statistical comparisons of the ISNR are realized with Welch's t-test (statistical significance: 0.05).

Algorithm 1 for local order-constraining deconvolution.

initialization: $n = 0$; $\boldsymbol{u}^{(0)} = \bar{\boldsymbol{u}}^{(0)} = \boldsymbol{z} \in \mathbb{R}^N$; $\boldsymbol{p}^{(0)} = 0 \in \mathbb{R}^N$; $\boldsymbol{q}^{(0)} = 0 \in \mathbb{R}^{dN}$; choose $\tau^{(0)} = \sigma^{(0)} = 0.9/\|(\boldsymbol{H}^T\boldsymbol{A}^T)^T\|_2$.
repeat
  $\boldsymbol{p}^{(n+1)} = \text{prox}_{\sigma^{(n)} F_1^\star}(\boldsymbol{p}^{(n)} + \sigma^{(n)}\boldsymbol{H}\bar{\boldsymbol{u}}^{(n)})$
  $\boldsymbol{q}^{(n+1)} = \text{prox}_{\sigma^{(n)} F_2^\star}(\boldsymbol{q}^{(n)} + \sigma^{(n)}\boldsymbol{A}\bar{\boldsymbol{u}}^{(n)})$
  $\boldsymbol{u}^{(n+1)} = \text{prox}_{\tau^{(n)} G}(\boldsymbol{u}^{(n)} - \tau^{(n)}(\boldsymbol{H}^*\boldsymbol{p}^{(n)} + \boldsymbol{A}^*\boldsymbol{q}^{(n)}))$
  $\bar{\boldsymbol{u}}^{(n+1)} = 2\boldsymbol{u}^{(n+1)} - \boldsymbol{u}^{(n)}$
until convergence of $\boldsymbol{u}$

VII. Results and discussion

Results of experiment 1 (see Fig. 2) show that, unlike theoretical bound, the empirical number of wrong inequalities in $\boldsymbol{A_H}$ increases slowly with S, even when $|\text{supp}(\boldsymbol{h})| > N/S$. This evolution is related to the shape of kernel $\boldsymbol{h}$. Asymetrical kernels (not shown here) give however results similar to bimodal kernels. In 2D, we could expect comparable results, considering separately each dimension as a signal in $\mathcal{C}(S)$. Figure 3 illustrates the results of experiment 2. Since LD cannot deal with the presence of noise, only results for $\sigma_n = 0$ are presented. As expected, this method leads to restored signals of poor quality, with Gibbs artifacts. In absence of noise, ROP provides significantly better results than other methods.

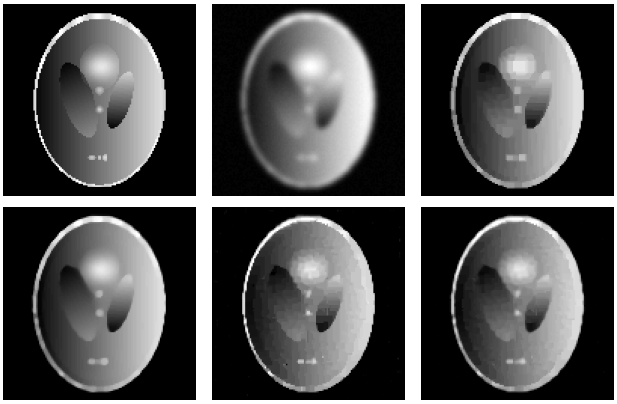

Fig. 4: Results on synthetic data. From left to right and top to bottom: original image, corrupted image (convolution with isotropic Gaussian kernel ($\sigma = 2$, $|\text{supp}| = 13 \times 13$) and corruption with additive Gaussian noise ($\sigma_n = 1$)), TV, TGV, OP and ROP restored images.

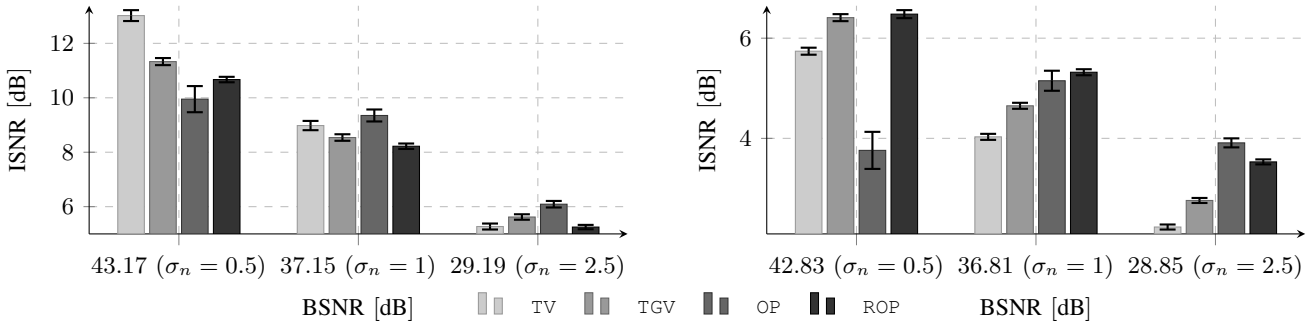

Fig. 5: Mean ISNR (100 realizations) of 2D images of size 128×128 restored with different methods (TV, TGV, OP and ROP). Original images are blurred by Gaussian unimodal kernel (left: $\sigma = 1.2$, right: $\sigma = 2$, $|\text{supp}| = 13 \times 13$) and corrupted by additive Gaussian noise ($\sigma_n = 0.5$, $\sigma_n = 1$ and $\sigma_n = 2.5$). Error bars represent the standard deviation of the sample.

Excepting results for ROP/TV when $(S = 7, \sigma_n = 0.005)$ and OP/ROP when $(S = 47, \sigma_n = 0.01)$, ISNR of the three methods are significantly different.

Results of experiment 3 are illustrated on Fig. 4 and Fig. 5. Except for TV/ROP when $(\sigma = 1.2, \sigma_n = 2.5)$, ISNR of images restored with OP and ROP methods are significantly different from each other and from TV and TGV ISNR. For small level of noise, TV, TGV and OP show a high increase in SNR compared to OP, as for the 1D results. For higher levels of noise, the proposed methods - OP or ROP - give the best results. Fig. 4 shows the restored images for $(\sigma = 2, \sigma_n = 1)$. ISNR are significantly different as well as their visual aspect: except for TGV which visually provides the best restoration, all methods produce images with staircasing artifacts. Unlike TV, the other methods, and in particular ROP, restore more accurately the small objects, which might improve the detectability of small cancerous lesions in PET images.

VIII. Conclusion

A deconvolution method equipped with a new regularization has been proposed. The driving idea is that convolution with blurring kernels has limited impact on the sign of derivatives. On the other hand, deconvolution tends to produce ringing artifacts that typically flip many derivative signs. To avoid such spurious oscillations without making strong assumptions about the image content, the proposed regularization enforces derivative signs in the deconvolved image, after estimating them from noisy measurements. In this way, the proposed method exploits information present in the observed image. Results validate the interest of the OP and ROP methods. ISNR values and visual aspect are comparable to those obtained with TV and TGV restoration and are even better for high levels of noise. Perspectives for future work aim at applying the proposed method to real PET images. For this purpose, the method has to be extended to three dimensions and Poisson noise, *e.g.*, through a variance stabilizing transform. Noise filtering and its desirable properties in this particular application will be investigated further, since they condition the correct estimation of derivative signs, *i.e.*, $\boldsymbol{A_H}$.

References

[1] J. A. Lee, "Segmentation of positron emission tomography images: some recommendations for target delineation in radiation oncology." *Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology*, vol. 96, no. 3, pp. 302–7, 2010.

[2] T. Shepherd, M. Teras, et al., "Comparative Study With New Accuracy Metrics for Target Volume Contouring in PET Image Guided Radiation Therapy," *IEEE Transactions on Medical Imaging*, vol. 31, no. 11, pp. 2006–2024, 2012.

[3] M. Soret, S. L. Bacharach, et al., "Partial-volume effect in PET tumor imaging." *Journal of nuclear medicine : official publication, Society of Nuclear Medicine*, vol. 48, no. 6, pp. 932–45, 2007.

[4] X. Geets, J. A. Lee, et al., "A gradient-based method for segmenting FDG-PET images: methodology and validation." *European journal of nuclear medicine and molecular imaging*, vol. 34, no. 9, pp. 1427–38, 2007.

[5] M. Wanet, J. A. Lee, et al., "Gradient-based delineation of the primary GTV on FDG-PET in non-small cell lung cancer: A comparison with threshold-based approaches, CT and surgical specimens," *Radiotherapy and Oncology*, vol. 98, no. 1, pp. 117–125, 2011.

[6] N. Boussion, C. Cheze Le Rest, et al., "Incorporation of wavelet-based denoising in iterative deconvolution for partial volume correction in whole-body PET imaging," *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 36, no. 7, pp. 1064–1075, 2009.

[7] S. Stute and C. Comtat, "Practical considerations for image-based PSF and blobs reconstruction in PET," *Physics in Medicine and Biology*, vol. 58, no. 11, pp. 3849–3870, 2013.

[8] A. Rahmim, J. Qi, et al., "Resolution modeling in PET imaging: theory, practice, benefits, and pitfalls." *Medical physics*, vol. 40, no. 6, 2013.

[9] Y. S. Lee, J. S. Kim, et al., "Performance measurement of PSF modeling reconstruction (True X) on Siemens Biograph TruePoint TrueV PET/CT," *Annals of Nuclear Medicine*, vol. 28, no. 4, pp. 340–348, 2014.

[10] K. Bredies, K. Kunisch, et al., "Total Generalized Variation," *SIAM Journal on Imaging Sciences*, vol. 3, no. 3, pp. 492–526, 2010.

[11] D. Bailey, D. Townsend, et al., *Positron emission tomography*, 2005.

[12] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge: Cambridge University Press, 2004.

[13] A. Chambolle and T. Pock, "A First-Order Primal-Dual Algorithm for Convex Problems with Applications to Imaging," *Journal of Mathematical Imaging and Vision*, vol. 40, no. 1, pp. 120–145, 2010.

[14] W. Hoeffding, "Probability inequalities for sums of bounded random variables," *Journal of the American statistical association*, 1963.

[15] C. W. Groetsch, "Differentiation of approximately specified functions," *The American Math. Monthly*, vol. 98, no. 9, pp. 847–850, 1991.

[16] M. Chiani, D. Dardari, et al., "New exponential bounds and approximations for the computation of error probability in fading channels," *IEEE Transactions on Wireless Communications*, vol. 2, no. 4, pp. 840–845, 2003.

[17] N. Parikh and S. Boyd, "Proximal algorithms," *Foundations and Trends in optimization*, vol. 1, no. 3, pp. 123–231, 2013.

[18] A. González, L. Jacques, et al., "Compressive optical deflectometric tomography: A constrained total-variation minimization approach," *Inverse Problems and Imaging*, vol. 8, no. 2, pp. 421–457, 2014.

[19] P. L. Combettes and J.-C. Pesquet, "Proximal Splitting Methods in Signal Processing," vol. 49, pp. 185–212, 2011.

[20] L. Condat, "A Direct Algorithm for 1-D Total Variation Denoising," *IEEE Signal Processing Letters*, vol. 20, no. 11, pp. 1054–1057, 2013.